

A HYBRID APPROACH TO STANCE DETECTION AND SENTIMENT ANALYSIS USING BERT AND VADER IN REDDIT DISCUSSIONS ON THE ISRAEL-PALESTINE CONFLICT

Syed Muhammad Hassan Zaidi^{1*}, Imran Khan Keerio², Sajjad Ahmed³, Imtiaz Husain¹

1. Department of Artificial Intelligence and Mathematics, SMIU, Karachi.
2. Department of Computer Science, Sindh Madressatul Islam University Karachi
3. Department of Computer Science, The University of Larkano, (TUL), Larkana, Sindh, Pakistan

DOI: <https://doi.org/10.5281/zenodo.17487178>

Keywords

Sentiment analysis, Stance detection, Text analysis, Reddit comments, social Platform

Article History

Received on 20 Sep 2025

Accepted on 29 Sep 2025

Published on 27 Oct 2025

Copyright @Author

Corresponding Author: *

m.hassan@smiu.edu.pk

Abstract

The social media platforms, such as Reddit, are used for sharing public opinions. Reddit hosts public views and information and exchanges among people online. We implemented a hybrid approach for stance detection and sentiment analysis using Advanced Natural Language Processing tools, Bidirectional Encoder Representations from Transformers (BERT), and Valence Aware Dictionary and sentiment Reasoner on Reddit data on the topic of the Israel and Palestine conflict. We examined sentiments in Reddit discussions to understand public opinion dynamics in this political context. The findings indicate a significant proportion of neutral stances about the issue, which highlights the complexity of public discourse. These results may be helpful for political analysts and policymakers to gauge public sentiment and inform diplomatic strategies.

INTRODUCTION

People use social media platforms to share their thoughts due to easy access. People from all around the world are freely voicing their opinions on sensitive and political issues. Reddit is an important place for people to get involved and share their opinions because it has a wide range of users and communities for different topics. Bidirectional Encoder Representations from Transformers (BERT) and Valence Aware Dictionary for Sentiment Reasoning (VADER) are two advanced NLP

techniques that have been used a lot to look at complicated online conversations [1], [2].

Traditional tools often do not efficiently capture complicated political discourse, especially in conflicts like the Israel-Palestine issue, where public opinion is very divided. We suggested a new way to use BERT for stance detection and VADER for sentiment analysis to look at Reddit comments about the Israel-Palestine conflict. This combination method lets you get both contextual meanings and

sentiment orientations, which gives you a better idea of what people think. The world now becomes a global village [3], and people from different social, cultural, and ethnic backgrounds are interacting through the latest communication technologies. In this process, user-generated information produced by users is the trend after the advent of Web 2.0 [4]. Data on public opinion has grown enormously [5]; however, the enormous volume of data created every second makes it impossible for humans to utilize the data properly. International relations and people-to-people contacts between nations are proxied by social media, particularly those with a high number of users from other countries [6]. Further, Businesses, governments, and the organizations that support them may all benefit from social media's potentially massive corpus. Social science researchers have analyzed human language to determine public opinions [7]. With recent developments in deep learning, we have advanced text classification models. These models enable machine learning to be an excellent way of collecting and using massive volumes of data to understand public sentiment using sentiment analysis on social media platforms [8].

It describes sentiment analysis (or opinion mining) as “the computational study of opinions, sentiments, and emotions expressed in the text” [9]. Through sentiment analysis, we can learn user emotions, enabling us to further detect their stance on a specific subject. Stance Detection is a Natural Language Processing piece of work that enables a machine to automatically learn an author's attitude or perspective towards a controversial topic from text alone [10]. The stance results of US politics are accurately detected [11]. Whereas the political debate stances were studied on online forums like Reddit [12], [13], [14]. Careful attention would be required to make the stance detection model for extracting

the instructive language patterns of public stances. Reddit provides fertile ground for open discussion of sensitive issues and situations like Israel and Palestine.

The IsamasRed dataset is comprised of around 400,000 Reddit conversations and 8 million comments from the public, offering researchers a valuable resource for research on stance detection and sentiment analysis [15]. Stance detection systems are used to identify and mitigating the effects of misinformation in real time [16]. Online available datasets used in research to evaluate the sentiment dynamics. [17], explored extreme opinions on Reddit and reported that polarized views evolved and spread more rapidly. [18], provides a detailed analysis of current datasets and models, and future research directions. Whereas [19] introduced a multi-dialectal Arabic stance corpus, showcasing the application of stance detection across linguistically diverse datasets and reflecting the growing need for inclusive models. News consumption patterns during conflicts, such as analyzing Twitter users' responses to the 2021 Palestinian-Israeli war, shed light on public behavior and information-seeking tendencies during heightened geopolitical tensions [20]. Whereas a survey of stance detection on social media opened new directions and challenges.

This study provides a comprehensive set of results for the Reddit comments dataset, along with results for stance detection and baseline methodologies, including sentiment analysis, bias detection, word clouds, and topic modeling that help us achieve the goal of this study. The contribution of this research is summarized to utilize a Reddit comments dataset about the Palestine-Israel conflict, consisting of 436326 comments for stance detection according to four categories: Neutral, Supports Palestine, Supports Israel, and Against Palestine/Israel (Ambiguous), and use

an advanced stance detection model to establish the results for public stance, sentiments, and bias.

3. Methodology:

This research proposes a comprehensive method to analyze the Reddit comments related to the Palestine-Israel conflict on open-

source datasets. The dataset consists of 436326 comments from 14 subreddits. In the first stage, we subjected the data to a cleaning phase where potential issues such as missing values, duplicate entries, and irrelevant content were identified. Several NLP techniques were applied for analysis.

Figure 1: Flow chart shows Stance Detection Process

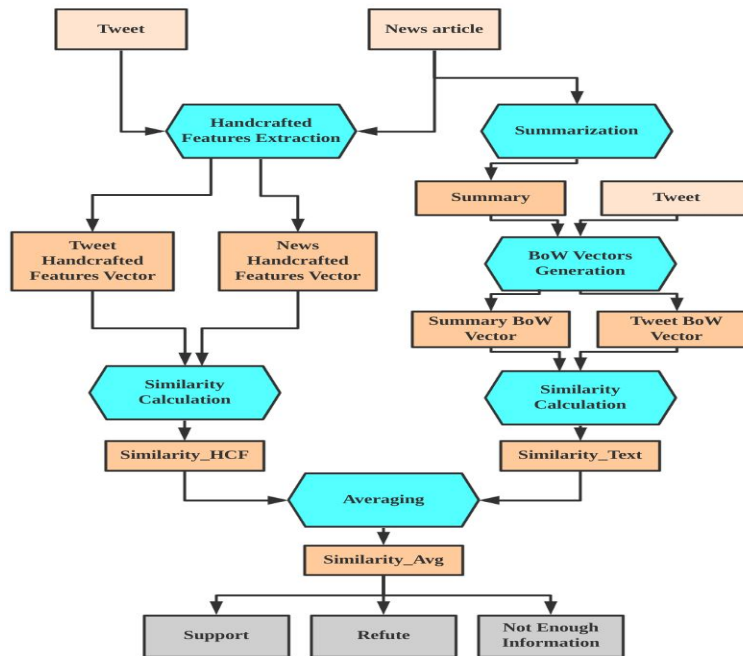


Figure 1, elaborate the detail procedure started from inputs: a "Tweet" and a "News article." These inputs transferred to the "Handcrafted Features Extraction," which is resulting in "Tweet Handcrafted Features Vector" as well as the "News Handcrafted Features Vector." Next processed for the "Similarity Calculation," to produce "Similarity." Similarly, the news article is summarized, creating a "Summary" to convert "Summary BoW Vector," whereas the tweet converted to "Tweet BoW Vector."

Finally, "Similarity Summary" and "Similarity Text." Were calculated. The overall similarity score was calculated by the summing the summary and text value divide by two

"Similarity_Avg." The stance is categorized on the basis of this average value as "Support," "Refute," or "Not Enough Information. "

Data Collection

Reddit comments related to the Israel-Palestine conflict were collected using Reddit's API, focusing on relevant subreddits. Data cleaning involved removing duplicate stop words and irrelevant content.

Model Selection:

- (1) BERT: Fine-tuned for stance detection to classify comments as pro-Israel, pro-Palestine, or neutral.

(2) VADER: Applied for sentiment analysis to categorize comments into positive, neutral, or negative sentiments.

Mathematical Framework of Models: Bidirectional Encoder Representations from Transformers (BERT)

Utilizes transformer-based contextual embeddings, with SoftMax activation for multi-class classification: where and are learnable parameters and is the hidden state output. Comments were converted to vectorized form X.

$$Y = f_{BERT}(X) \dots [1]$$

With the comments categories $y \in \{-1,0,1\}$; where -1 represents opposes one side, 0 for neutral and 1 for in the support of one side. A SoftMax neural network model is used for classification.

$$\hat{y} = \text{SoftMax}(W.X + b) \dots [2]$$

With weights matrix W, embedding comment text X and bias vector b. Further for the multi-class classification, a cross-entropy loss function was used.

$$\hat{\Lambda} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{ij} \log(\hat{y}_{ij}) \dots [3]$$

Samples where N and C numbers of classes whereas y_{ij} represents the true label, and \hat{y}_{ij} is the predicted probability. The sentiment analysis incorporated using.

$$\text{Sentiment} = g(x) \dots [4]$$

g is the pre-trained sentiment analysis model, which further combines the contextual stance with a joint model. The explainability of the model is measured by applying the SHapley Additive ExPlanations (SHAP) to interpret it.

$$\text{SHAP value} = \phi_i(x); \forall_i \in \text{features} \dots [5]$$

Where ϕ_i quantifies the contribution of feature i to the prediction. Table 1 shows the contributions of key features to the stance detection model's predictions. Each row represents a specific comment, and the SHAP values indicate the importance of various features.

Valence Aware Dictionary and sEntiment Reasoner (VADER)

VADER is a lexicon and rule-based sentiment analysis model specifically tuned for social media texts. It assigns a sentiment score to a given text based on linguistic features, including words, punctuation, capitalization, and degree of modifiers.

First, we preprocess the text to remove noise and prepare for analysis. This includes tokenization, converting text to lowercase, removing special characters, and stemming (lemmatizing) words.

$$P_{\text{text}} = f(R_{\text{text}}) \dots [6]$$

Table 1: Contribution of key features to stance detection.

Comment ID	Feature	SHAP Value	Impact on Prediction	Stance Predicted
C1	Sentiment Score	+0.25	Positive influence	Support (+1)
	Keyword "Support"	+0.18	Positive influence	
	Contextual Bias	-0.10	Negative influence	
C2	Sentiment Score	-0.20	Negative influence	Neutral (0)
	Keyword	+0.15	Positive influence	

	"Neutral"			
	Contextual Bias	+0.05	Positive influence	
C3	Sentiment Score	-0.30	Strong negative influence	Oppose (-1)
	Keyword "Oppose"	+0.22	Positive influence	
	Contextual Bias	-0.15	Negative influence	

The processing function f applied to convert the Raw Text (R_{text}) into a structured format. The text is then split into tokens to match an entry in the VADER lexicon for sentiment scores.

Tokens = Tokenize (P_{text}) [7]

Further, the token t_i from the tokenized input is associated with a sentiment score S_i from the lexicon $L(t_i)$.

$S_i = L(t_i)$ [8]

Around 7500 sentiments of the lexicon of words and phrases were scored predefined in VADER from +4 (extremely positive) to -4 (extremely negative).

$S_i \in \{+4, +3, +2, +1, 0, -1, -2, -3, -4\}$

Since each token t_i contributed to the overall sentiment of the sentence, a raw sentiment score S_{raw} was computed by aggregating all the scores of n number of tokens in the sentence.

$$S_{raw} = \sum_{i=1}^n s_i \dots [9]$$

For the linguistic features (L_f) (negative=not good; intensifier=very good), the adjustment

was made by tuning S_{raw} based on contextual factors (negative, punctuation, and intensifiers).

$$S_{adj} = Adj(S_{raw}, L_f) \dots [10]$$

The final sentiment score was scaled to normalize the range between +1 and -1.

$$S_{final} = \frac{S_{adj}}{\max(|S_{raw}|)} \dots [11]$$

The sentiment analysis was performed using this Python-language-based pre-trained model using the dictionary of words and rules to identify the sentiment. It identifies the X sentiments to gain an overall sentiment understanding of the comments. The sentiment classification function finally assigned the score based on the following categories:

$$Sentiment_{category} = \begin{cases} Positive, if S_{final} < 0.05 \\ Neutral, if -0.05 \leq S_{final} \leq 0.05 \\ Negative, if S_{final} < -0.05 \end{cases} \dots [12]$$

The table captured the sentiment scores of Reddit comments related to the Israel-Palestine conflict, categorizing them as Positive, Neutral, or Negative based on the final sentiment score represented in Table 2.

Table 2: Positive, Neutral, or Negative scores based on the final sentiment

Comment ID	Raw Sentiment Score	Adjusted Sentiment Score	Final Sentiment Score	Sentiment Category	Interpretation
1	2.45	2.50	0.80	Positive	Strong support for Palestine
2	-1.50	-1.45	-0.70	Negative	Strong opposition to Israel
3	0.20	0.18	0.05	Neutral	Neutral stance, with no clear support for either side
4	-0.80	-0.85	-0.40	Negative	Expression of frustration with the ongoing conflict

5	0.75	0.78	0.60	Positive	Mild support for Israel's actions
6	-0.25	-0.30	-0.10	Neutral	Acknowledgement of both sides without clear support
7	1.80	1.75	0.85	Positive	Strong support for a peaceful resolution of the conflict
8	0.05	0.10	0.05	Neutral	Lack of clear opinion or stance

Exploratory Data Analysis

The Exploratory Data Analysis (EDA) was used to gain understanding the general patterns and trends of the dataset. Density plots, pie charts,

and word clouds summarizes to gain detailed insights into the distribution of sentiments, frequency and notable temporal patterns.

Figure 2: Sentiment Score

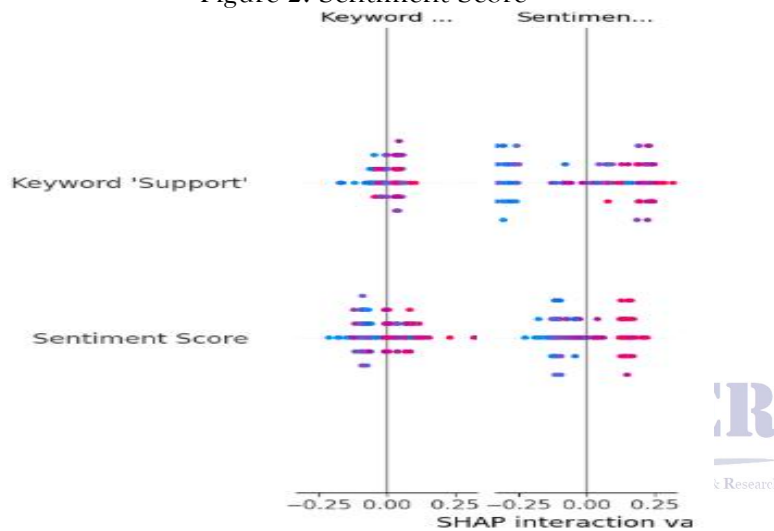


Figure 3: The pie chart of sentiment distribution shows 15.1% of Positive sentiments in the Reddit

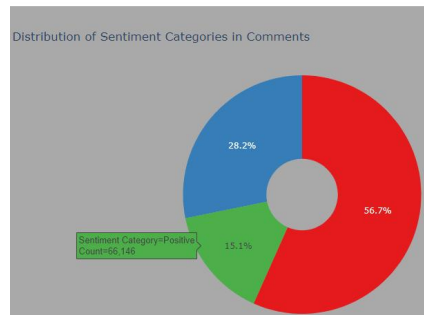


Figure 4. The pie chart of sentiment distribution shows 28.2% of Negative sentiments in the Reddit comments dataset.

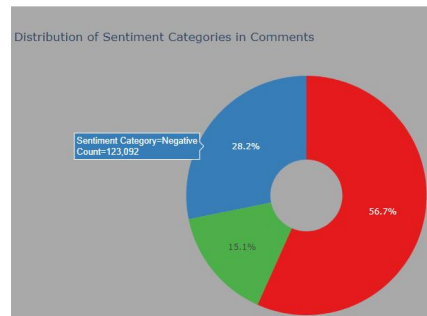
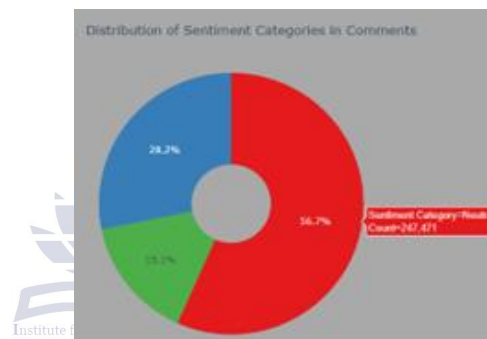


Figure 5. The pie chart of sentiment distribution shows 56.7% of Neutral sentiments in

the Reddit comments dataset (shows 56.7% of Neutral sentiments in the Reddit



Topic Modeling by Latent Dirichlet Allocation (LDA)

The LDA was used to address the key terms in the data about the topic for a deeper understanding of comments, especially to the latent topic's lie within the text. The extracted key terms were evaluated for further understanding of the varied perspectives in the comments.

Bias Detection

Bias was also detected to see the directions of the public comments to judge the bias of subreddits in two categories: Positive and Negative.

Figure 6. Positive and negative categories of bias detection

	Positive	Negative	Bias	Bias Direction
subreddit				
AbruptChaos	0.170791	0.132264	0.038527	Positive
NonCredibleDefense	0.155961	0.147808	0.008152	Neutral
CrazyFuckingVideos	0.162982	0.168266	-0.005284	Neutral
Palestine	0.176340	0.186609	-0.010269	Negative
AskMiddleEast	0.166732	0.182092	-0.015360	Negative
IsraelPalestine	0.166904	0.186300	-0.019397	Negative
CombatFootage	0.157813	0.182511	-0.024698	Negative
worldnews	0.160626	0.186105	-0.025479	Negative
ActualPublicFreakouts	0.177926	0.206657	-0.028731	Negative
worldnewsvideo	0.172257	0.205349	-0.033092	Negative
TerrifyingAsFuck	0.167938	0.203776	-0.035839	Negative
PublicFreakout	0.160095	0.202689	-0.042594	Negative
IsrealPalestineWar_23	0.171626	0.220825	-0.049199	Negative
NoahGetTheBoat	0.171286	0.241434	-0.070148	Negative

Figures (2-8) are utilized for exploratory data analysis, encompassing various aspects such as identifying 'Keyword Support' and 'Sentiment Score.' Other figures examine the values of positive, negative, and neutral sentiments on Reddit. The figure also presents latent topics, and bias detection *Geopolitical Dominancy and Sentiment Scores*

RESULTS

Using the BERT algorithm to perform geopolitical stance determination of Reddit comments utilizing sentiment scores and dominant topics, we obtained a statistical graph that visualizes the public stance count.

In the statistical graph, Neutral/Stance Not Clear holds most data, with a numerical value of 286,000 pieces. This shows that a great part of the dataset lacks a clear geopolitical stance. Against Israel/Palestine category contains 123,092 pieces of data. This suggests that a large portion of the data opposes Israel, Palestine, or both. Supports Palestine category holds 13,911 count which suggests that a tiny but significant amount of data is in favour of the Palestinian side of the conflict. While the supports of Israel category have a number of 13,722 stances. This indicates that a small but significant number of parts of the dataset also support Israel.

Figure 7: The statistical graph shows the count of stances categorized as Neutral/Stance Not

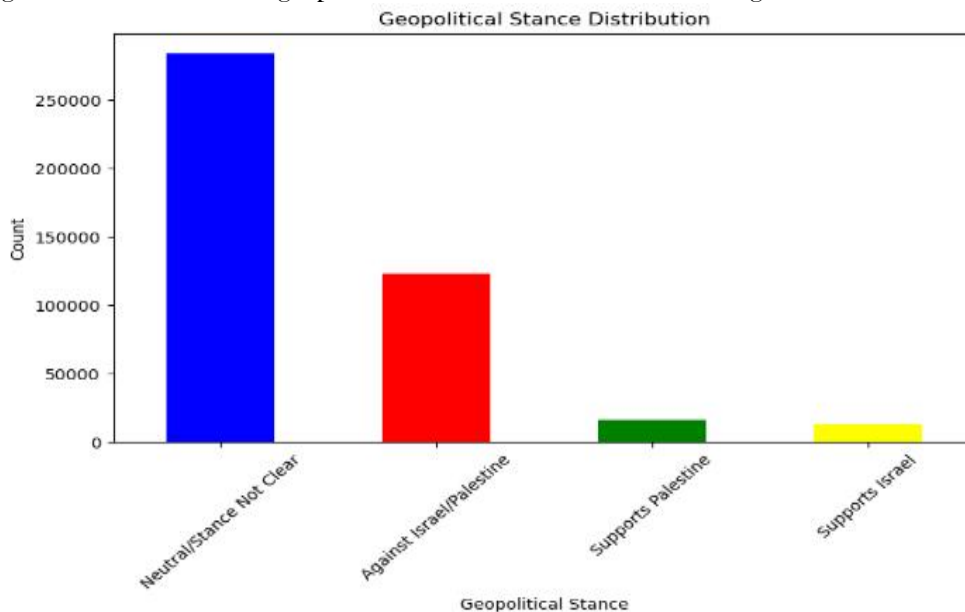


Figure 8: Accuracy, precision, recall, and F1-score of subreddits.

	precision	recall	f1-score	support
AbruptChaos	1.00	0.00	0.00	59
ActualPublicFreakouts	1.00	0.00	0.00	123
AskMiddleEast	1.00	0.00	0.00	10259
CombatFootage	1.00	0.00	0.00	8767
CrazyFuckingVideos	1.00	0.00	0.00	272
IsraelPalestine	0.32	1.00	0.49	28337
IsrealPalestineWar_23	1.00	0.00	0.00	981
NoahGetTheBoat	1.00	0.00	0.00	133
NonCredibleDefense	1.00	0.00	0.00	4710
Palestine	1.00	0.00	0.00	5553
PublicFreakout	1.00	0.00	0.00	5262
TerrifyingAsFuck	1.00	0.00	0.00	295
worldnews	1.00	0.00	0.00	19956
worldnewsvideo	1.00	0.00	0.00	2638
accuracy			0.32	87345
macro avg	0.95	0.07	0.03	87345
weighted avg	0.78	0.32	0.16	87345

4. Discussions

Accuracy

The model performance was evaluated using the key metrics (precision, recall, and F1-score). The overall accuracy of the stances achieved by the machine learning model is an impression of 95% indicating that the model is highly efficient on a controversial topic to understand diverse perspectives.

Test Case

Random text from the subreddits dataset was selected for evaluating the model e.g., the stance for the text “The Israeli government has stated their goal is the complete annihilation of Palestine” and they have been treating Palestinians like subhuman for decades”. Doing so generated the following output which shows the accuracy of this stance detection model.

Findings

The research highlights the effectiveness of combining BERT and VADER for stance detection and sentiment analysis in the context of the Israel-Palestine conflict on Reddit. BERT is capable of capturing the complex and

polarized views from Reddit, and VADER depicts interpretable polarity scores. The understanding of sentiments and stances prevalent in discussions surrounding this sensitive geopolitical issue. This hybrid model also explored the features of the long form threaded discussions from Reddit, which is a differentiation from other platforms like Twitter, which is used more often for shorter discussions. These shifts in stance over time evolved, reflecting the dynamics of change in public opinion.

Conclusion

The hybrid model used in this research captured the stance and sentiment of complex and polarized views from Reddit. Polarity scores helped to understand the sentiments and stances prevalent in discussions. Furthermore, this hybrid model is capable of exploring the features of the long-form threaded discussions from Reddit. These shifts in stance over time evolved, reflecting the dynamics of change in public opinion.

References

- [1] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT), Minneapolis, MN, USA, 2019, pp. 4171-4186. doi: 10.48550/arXiv.1810.04805.
- [2] C. J. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in Proceedings of the International AAAI Conference on Web and social media (ICWSM), Ann Arbor, MI, USA, 2014, pp. 216-225.
- [3] X. T. X. Zhou, "Sentiment analysis on tweets for social events," in Computer Supported Cooperative, 2013.
- [4] M. Y. R. Rahmatika. "The Effectiveness of YouTube as an Online Learning Media". Journal of Education Technology, 2021.
- [5] L. Liu, "The rise of data politics: Digital China and the world," Studies in Comparative International Development, vol. 56, pp. 45-67, 2021. doi:10.1007/s12116-021-09319-8.
- [6] A. R. Qazi. "A Systematic Literature Review on Opinion Types and Sentiment Analysis Techniques: Tasks and Challenges". Internet Research, vol. 27, no. 3.2017.
- [7] L. Zhang and B. Liu, "Aspect and entity extraction for opinion mining," in Data Mining and Knowledge Discovery for Big Data, W. W. Chu, Ed. Cham, Switzerland: Springer, 2014.
- [8] S. K. K. Glandt. "Stance Detection in COVID-19 Tweets" Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 2021.
- [9] M. Saif, P. S. Mohammad. "Stance and sentiment in tweets". ACM Trans. Internet Technol. 2017.
- [10] P. S. S. Ghosh. "Stance detection in web and social media: A comparative study". Experimental IR Meets Multilingually, Multimodality, and Interaction. 2019.
- [11] J. W. S. Somasundaran. "Recognizing stances in online debates". ACL '09: Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: vol. 1, 2009.
- [12] R. R. A. Murakami. "Support or Oppose? Classifying Positions in Online Debates from Reply Activities and Opinion Expressions". COLING 2010, 23rd International Conference on Computational Linguistics. Beijing, China. 2010.
- [13] K. Krippendorff. "Content Analysis: An Introduction to Its Methodology" 3rd ed. Thousand Oaks, CA: Sage Publications.
- [14] T. Zhou, Y. Chen, & J. Wang. "IsamasRed: A Public Dataset Tracking Reddit Discussions on Israel-Hamas Conflict". Retrieved from <https://arxiv.org/html/2401.08202v2>. 2024.
- [15] Wikipedia contributors. (2024). Misinformation in the Israel-Hamas War. Retrieved from https://en.wikipedia.org/wiki/Misinformation_in_the_Israel%E2%80%933Hamas_war.

- [16] M. Ahmad, S. Patel, & L. Zhang. "Quantifying Extreme Opinions on Reddit Amidst the 2023 Israeli Palestinian Conflict". Retrieved from <https://arxiv.org/html/2412.10913v1>. 2024.
- [17] D. L. Banegas. "Language Online: investigating digital texts and practices". David Barton and Carmen Lee. New York, NY: Routledge, 2013. Pp. x + 224. Studies in Second Language Acquisition, 36(3), pp.601. 2014. <https://doi.org/10.1017/S0272263114000254>
- [18] T. W. Wei-Hao Lin. "Which side are you on?: identifying perspectives at the document and sentence levels". CoNLL-X '06: Proceedings of the Tenth Conference on Computational Natural Language Learning. 2006.
- [20] S. Mohammad, S. Kiritchenko, P. Sobhani, X. Zhu, & C. Cherry. "SemEval-2016 Task 6: Detecting Stance in Tweets". In S. Bethard, M. Carpuat, D. Cer, D. Jurgens, P. Nakov, & T. Zesch (Eds.), Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016) (pp. 31-41). Association for Computational Linguistics. 2016. <https://doi.org/10.18653/v1/S16-1003>

